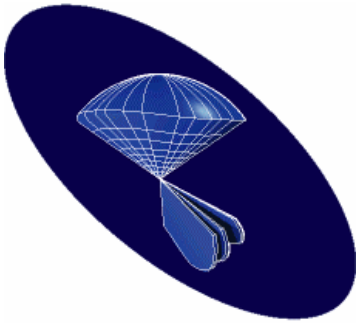

SDSS-II Development Projects

Bill Boroski

SDSS-II First Year Review
National Science Foundation
August 7, 2006

Areas of Development

- **SEGUE Survey**
- **Supernova Survey**
- **Photometric Calibration**
- **Data Acquisition System Upgrade**

SEGUE Development Work (1)

- **Target Selection**
  - Algorithm for high-latitude (|b|>20) stellar objects done
  - Algorithm for low-latitude objects has been tested on the sky; results still being analyzed.
    - *Simpler algorithm focusing on blue stars, K-giants, high proper motion objects.*

SEGUE Development Work (2)

- **Spectro Data Reduction Pipeline (Spectro v5)**
  - Several new outputs
    - Extracted spectra for a single CCD of a single 15 minute exposure, before red and blue halves, and multiple exposures, are combined.
    - Sky flux subtracted from each fiber.
    - New header keywords
      - Median and RMS spectrophotometry minus photo mag for standard stars and main galaxies in g, r, and i-bands
      - Exposure number for the "best" exposure (all 4 cameras)
      - Identifier for each individual exposure + CCD used in the exposure combine
    - New fields in the spPlate and spCFrame files:
      - ‘SCI_EXPTIME‘, 'CALIBFLUX ‘, CALIBFLUX_IVAR‘, and 'SFD_EBV'
  - Testing and analysis is underway
    - Improved scattered light handling
    - Zero points of the radial velocity templates (ELODIE velocities are off slightly – currently being studied)
  - Plan is to reprocess all spectro data using the new pipeline once it’s declared production-ready
  - New reductions will be available in DR6

SEGUE Development Work (3)

- **Stellar Atmosphere Pipeline**
  - Used to estimate stellar parameters (Teff, log g, [Fe/H]) based on R=2000 spectroscopy and ugriz photometry.
  - Incorporates a number of independent methods (obtained from different calibrations) for each parameter, which are then suitably averaged in the final estimation process.
  - Version 1.0 was recently completed
    - Run by the development team and an initial set of stellar parameters distributed to the Collaboration for review - Testing and validation by the production team should begin soon.
  - Remaining work includes
    - Finishing observations of first-pass calibration stars now being obtained with HET; when the data are reduced and analyzed, we anticipate having over 100 high-res spectra of SEGUE stars with which to evaluate Spectro v5.
    - Continue work to extend the reach of the current pipeline to include stars with Teff > 10,000K.
    - Refine methods for estimating stellar parameters based solely on ugriz photometry.
    - Develop and test an artificial neural network approach for parameter estimation from spectroscopy alone.

SEGUE Development Work (4)

- **Calibration of Filter Systems, Photometry**
  - Refine transforms: ugriz <->UBVRI <-> u'g'r'i'z'
  - Work being carried out at several institutions, with scientific goal of obtaining photometric metalicities

- **Photometric Pipeline**
  - Development work involves modifying Photo to process data taken in crowded fields
  - The code has been adapted to handle crowded-field photometry on multiple-band data at low latitudes.
  - Testing is underway by the developers using test fields and reductions.
  - Remaining work includes
    - Packaging the code for inclusion into production data processing operations
    - Delivering the code to production for testing and validation
    - Processing existing SEGUE data using the new pipeline; and distributing the results to the Collaboration for testing and analysis
    - Incorporating the new pipeline into the production process
    - Updating the CAS data model to accept the additional data outputs

SEGUE Development Work (5)

- **Quality Assessment Tools**
  - Some work ongoing at a low level to develop SEGUE-specific QA tests
  - Data quality currently being monitored using existing imaging and spectro QA tools
  - Quality also monitored by scientists looking at various data types and by developers in the process of code development & testing

- **Data Processing System at Princeton**
  - SEGUE spectroscopic reductions will eventually occur at Princeton
  - Computing hardware has been purchased, installed and commissioned
  - Additional hardware (with substantial disk space) will be required to support imaging data processing on crowded fields.

- **Data Distribution Upgrades**
  - No significant progress to date; efforts were focused on completing DR5
  - SEGUE data have been distributed using existing DAS and CAS
  - Approved data model changes for DR6 geared to accommodate various SEGUE parameters
  - Modifying database code to accommodate these changes is currently underway

Supernova Survey Development (1)

- **SN Compute Cluster**
  - A rack of (10) dual-processor machines was assembled and installed at APO to perform near-real-time reductions of SN data
  - DC-powered computers were chosen to reduce heat load
  - Data are spooled directly from the DA file servers throughout the night
  - Artificial supernovae were used to monitor the performance of the data system, checking detection efficiency and photometric accuracy as functions of magnitude for each run.
  - System worked very reliably throughout the fall 2005 observing season.

- **APO Computer Room Upgrade**
  - Three existing Uninterruptable Power Supplies (UPS) were replaced with a single, more efficient 20 KVA unit.
  - New unit was relocated outside of the computer room to further reduce heat loading
  - Significantly less costly, and less-disruptive solution than upgrading plant cooling system

Supernova Survey Development (2)

- **Frame Subtraction Pipeline**
  - Modified to use more astrometric, PSF and mask information from Photo to reduce the number of processing steps.
  - Incorporated capability to handle co-added Stripe 82 template data, intended to improve the S/N for supernova detection.
  - Implemented improved diagnostics of subtraction performance for QA
  - Implemented improved remapping algorithm for more reliable subtractions.
  - Improved PSF convolution and object-finding algorithms.
  - Added on-mountain subtraction in i-band (in addition to g and r-band subtractions that were there for the 2004 test run)

- **Stripe 82 Database**
  - Created a simple MySQL database containing the SN imaging runs obtained on Stripe 82
  - Created a veto catalog of variable sources and a catalog of calibration stars.
  - Over time, this will likely be merged with the Runs DB

Supernova Survey Development (3)

- **Spectroscopic Target Selection**
  - Developed a web-based interface using gif images to improve the process of human inspection of supernova candidates.
  - Implemented photometric redshift information for supernova candidate host galaxies to aid in target selection.
  - Developed an automated target selection algorithm using early-epoch photometry data and fits to redshifted, multi-band template light-curves of different supernova types
  - Augmented automated algorithm with further information about host galaxy (and likely host galaxy extinction) from the SDSS database.

- **SN Data Distribution**
  - A web site for distributing SN data was developed and made available to the public in January 2006.
  - Provides an overview of the SN Survey, a description of the data products, and instructions for accessing the data.

Photometric Calibration

- **Objective:** *to improve the relative photometric calibration of the SDSS imaging data from of order 2% to of order ~1%.*

- **To date, we have successfully:**
  - Run all Apache Wheel data through the Apache Wheel pipeline
  - Developed code to incorporate these data into the ubercalibration pipeline
  - Developed code to sweep through the SDSS imaging data, in order to create a list of stars based on magnitude limits set for each of the five filters
  - Developed a full survey simulator to test the calibrations
  - Ran the selected set of imaging data through the ubercal pipeline
  - Result is ~1% relative calibration errors across 8500 deg^2 in griz and ~2% for the u band.
  - A paper has been written by Nikhil Padmanabhan, et al, describing the process and results in detail. Currently in draft form; distributed internally for review.

Photometric Calibration (2)

- **Remaining Work**
  - Ubercal generates its own flat field vectors, which are different from the flat fields used in the production reduction; effort is required to determine if/how to apply corrections.
  - Ubercal only works on bright isolated stars; effort is required to determine if corrections can accurately be applied to other objects
  - Exporting ubercal outputs for inclusion in production data processing operations;
    - Still need to formulate a plan for how ubercal outputs will be applied to the factory reductions, or how ubercal might be incorporated into factory operations
    - One option is to add a table to the CAS containing the ubercal corrections; users can choose whether or not to apply.

DA Upgrade

- **Objective:** *to address obsolescence and reliability issues; eliminate writing data to tape; and provide faster access to data.*

- **The new system was put into place during the 2005 summer shutdown**
  - Obsolete SGI Challenge computers were replaced with Linux boxes
  - Obsolete MVME167 single-board computers were replaced with MVME5500 PowerPC computers.
  - Fileservers with removable drive bays were implemented. Data are now written directly to disk; each server has capacity for (9) nights of imaging data.
  - Modifications were made to several software packages (including porting code to run on Linux)
  - Added new features (e.g., checksum calculations)

- **Modifications at the receiving end**
  - Developed automated scripts that look for new data, initiate data transfer over the internet, and package the data for archiving and data processing
  - Created web page that reports the status of transferred data and informs the observers of data that can be deleted from the APO file servers

DA Upgrade (2)

- **The Result**
  - The new system has proven to be more reliable and robust
  - Data taken with new DA are of the same quality as those taken prior to the upgrade.
  - Upgrade was completed on time: first science spectra collected on August 29 and first imaging data on Sept 2.
  - Although there were numerous communication problems and minor bugs to work out, very minimal science time was lost to implementation.